Sandro Marini[1], MD; Katherine Crawford[1], BS; Andrea Morotti[2], MD;

Myung J. Lee[3], BA; Alessandro Pezzini[4], MD; Charles J. Moomaw[5], PhD;

Matthew L. Flaherty[5], MD; Joan Montaner[6,7,8], MD, PhD; Jaume Roquer[9], MD, PhD;

Jordi Jimenez-Conde[9], MD, PhD; Eva Giralt-Steinhauer[9], MD, PhD;

Roberto Elosua[9], MD, PhD; Elisa Cuadrado-Godia[9], MD, Ph;

Carolina Soriano-Tarraga[9], PhD, BSc; Agnieszka Slowik[10], MD, PhD;

Jeremiasz M. Jagiella[10], MD, PhD; Joanna Pera[10], MD; Andrzej Urbanik[10], MD, PhD;

Alexander Pichler[10], MD; Björn M. Hansen[11], MD; Jacob L. McCauley[12], PhD;

David L. Tirschwell[13], MD, MSc; Magdy Selim[14], MD, PhD;

Devin L. Brown[15], MD, MS; Scott L. Silliman[16], MD; Bradford B. Worrall[17], MD, MSc;

James F. Meschia[18], MD; Chelsea S. Kidwell[19], MD; Fernando D. Testai[20], MD;

Steven J. Kittner[21], MD, MPH; Helena Schmidt[22], MD; Christian Enzinger[22], MD;

Ian J. Deary[23], FBA, FRSE, FMedSci; Kristiina Rannikmae[24], MD, PhD;

Neshika Samarasekera[24], PhD, MRCP; Rustam Al-Shahi Salman[24], MA, PhD, FRCP;

Catherine L. Sudlow[24], BMBCh, MSc, DPhil, FRCPE;

Catharina J.M. Klijn[26], MD, PhD; Koen M. van Nieuwenhuizen[26], MD;

Israel Fernandez-Cadenas[9], PhD; Pilar Delgado[6], MD, PhD; Bo Norrving[11], MD;

Arne Lindgren[11], MD; Joshua N. Goldstein[3], MD, PhD; Anand Viswanathan[3], MD, PhD;

Steven M. Greenberg[3], MD, PhD; Guido J. Falcone[27], MD, ScD, MPH;

Alessandro Biffi[28], MD; Carl D. Langefeld[29], PhD; Daniel Woo[5], MD;

Jonathan Rosand[1,3,30], MD, MSc and Christopher D. Anderson[1,3,30], MD, MMSc

on behalf of the International Stroke Genetics Consortium

1 Center for Genomic Medicine, Massachusetts General Hospital, Boston, MA, USA

2 Stroke Unit, IRCCS Mondino Foundation, Pavia, Italy

3 Department of Neurology, Massachusetts General Hospital, Boston, MA, USA

4 Department of Clinical and Experimental Sciences, Neurology Clinic, University of Brescia, Brescia, Italy

5 Department of Neurology and Rehabilitation Medicine, University of Cincinnati College of Medicine, Cincinnati, OH, USA

6 Neurovascular Research Laboratory and Neurovascular Unit, Institut de Recerca, Hospital Vall d'Hebron, Universitat Autonoma de Barcelona, Barcelona, Spain

7 Institute de Biomedicine of Seville, IBiS/Hospital Universitario Virgen del Rocío/CSIC/University of Seville, Seville, Spain;

8 Department of Neurology, Hospital Universitario Virgen Macarena, Seville, Spain

9 Department of Neurology, Neurovascular Research Unit, Institut Hospital del Mar d'Investigacions Mèdiques, Universitat Autonoma de Barcelona, Barcelona, Spain

10 Department of Neurology, Jagiellonian University Medical College, Krakow, Poland

11 Department of Clinical Sciences Lund, Neurology, Lund University, Lund, Sweden and Department of Neurology and Rehabilitation Medicine, Skåne University Hospital, Lund, Sweden

12 John P. Hussman Institute for Human Genomics, University of Miami, Miller School of Medicine, Miami, FL 33136, USA

13 Stroke Center, Harborview Medical Center, University of Washington, Seattle, WA, USA

14 Department of Neurology, Stroke Division, Beth Israel Deaconess Medical Center, Boston, MA, USA

15 Cardiovascular Center, University of Michigan, Ann Arbor, MI, USA

16 Department of Neurology, University of Florida College of Medicine, Jacksonville, FL, USA

17 Department of Neurology and Public Health Sciences, University of Virginia Health System, Charlottesville, VA, USA

18 Department of Neurology, Mayo Clinic, Jacksonville, FL, USA

19 Department of Neurology, University of Arizona, Tucson, AZ, USA

20 Department of Neurology and Rehabilitation, University of Illinois College of Medicine, Chicago, IL, US

21 Department of Neurology, Baltimore Veterans Administration Medical Center and University of Maryland School of Medicine, Baltimore, MD, USA

22 Department of Neurology, Medical University of Graz, Graz, Austria

23 Centre for Cognitive Ageing and Cognitive Epidemiology, University of Edinburgh, Edinburgh, United Kingdom

24 Centre for Medical Informatics, Usher Institute of Population Health Sciences and Informatics, University of Edinburgh, United Kingdom

25 Centre for Clinical Brain Sciences, University of Edinburgh, United Kingdom

26 Department of Neurology, Radboud University Medical Centre; Donders Institute for Brain, Cognition and Behaviour, Nijmegen, and Department of Neurology and Neurosurgery, Brain Center Rudolf Magnus, University Medical Center Utrecht, Utrecht, the Netherlands

27 Division of Neurocritical Care and Emergency Neurology, Department of Neurology, Yale University School of Medicine, New Haven, Connecticut, USA and Center for Neuroepidemiology and Clinical Neurological Research, Yale School of Medicine, Yale University, New Haven, CT, USA

28 Division of Behavioral Neurology, Massachusetts General Hospital, Boston, MA, USA

29 Center for Public Health Genomics and Department of Biostatistical Sciences, Wake Forest University, Winston-Salem, NC, USA

30 Program in Medical and Population Genetics, Broad Institute, Cambridge, MA, USA

Correspondence:

Christopher D. Anderson MD, MMSc

185 Cambridge Street CPZN 6818

Boston, MA 02114

Phone: +1-617-726-4369

e-mail: cdanderson@partners.org

- **Date of the revision: 10/23/2018**

- **Title: Apolipoprotein E Alleles and Risk of Intracerebral Hemorrhage: a Trans-Ethnic Meta-Analysis**

- **Title characters count: 75**

- **Running head:** Trans-ethnic Meta-Analysis of APOE in ICH

- **Running head characters count**: 42

- **Manuscript words**: 3840

- number of figures 4

- number of tables 1

- Supplemental: 4 tables, 1 figure

Key Points

Question: What is the effect of history of hypertension and Apolipoprotein E (APOE) on intracerebral hemorrhage (ICH) risk in subjects stratified by self-reported race/ethnicity?

Findings: In this case control study that included 13,124 adults, having a copy of *APOE* ε4 increased the risk for lobar ICH only in whites, but after propensity score-matching for hypertension burden, Hispanics subjects showed the same effect of *APOE* ε4.

Meaning: *APOE* ε4 is confirmed to be a risk factor for lobar ICH. Its effect is present in non-white populations but is masked by differential hypertension burden.

Abstract

Importance

Genetic studies of intracerebral hemorrhage (ICH) have focused mainly on white participants. Genetic risk may vary or could be concealed by differing non-genetic co-exposures in non-white populations. Trans-ethnic analysis of risk may clarify the role of genetics in ICH risk across populations.

Objective

We sought to determine whether established differences in ICH risk by race and ethnicity could be due to variability in the effects of Apolipoprotein E (*APOE*) epsilon (ε) alleles, the most potent genetic risk factor for ICH.

Design, Setting and Participants

. We meta-analyzed the effects of *APOE* allele status on ICH risk, applying a two-stage clustering approach based on race/ethnicity and contributing study. A propensity score analysis was used to model the influence of *APOE* against the burden of hypertension across races/ethnicities. Primary ICH cases and controls were collected from hospital- and population-based studies from US and European sites within the International Stroke Genetic Consortium, enrolled from 1999 to 2017.Secondary causes of ICH were excluded from enrollment. Controls were regionally matched within each participating study.

Clinical variables were systematically obtained from structured interview within each site. *APOE* genotype was centrally determined for all studies.

Results

13,124 subjects (54.5% male, median age 66 (18-100) years) were included. In whites, *APOE* ε2 (odds ratio (OR)=1.49, 95% confidence interval (CI)=1.24-1.80, p<0.001) and *APOE* ε4 (OR=1.51, 95% CI=1.23-1.85, p<0.001) were associated with lobar ICH risk, however within self-identified Hispanics and blacks, no associations were found. After propensity score-matching for hypertension burden, *APOE* ε4 was associated with lobar ICH risk among Hispanics (OR=1.14, 95% CI=1.03-1.28, p=0.01), but not in blacks (OR=1.02, 95% CI=0.98-1.07, p=0.251). *APOE* ε2 and ε4 did not show an effect on non-lobar ICH risk in any race/ethnicity.

Conclusions and Relevance

APOE ε4 and ε2 alleles affect lobar ICH risk variably by race and ethnicity. Associations are confirmed in whites but can be shown in Hispanics only when the excess burden of hypertension is propensity score-matched. Further studies are needed to explore interactions between *APOE* alleles and environmental exposures that vary by race and ethnicity in representative populations at risk for ICH.

INTRODUCTION

Spontaneous intracerebral hemorrhage (ICH) is the most severe form of stroke. In the US alone, 160,000 people suffer from ICH each year with a case fatality rate of 54% at 1 year[1]. ICH prevalence has increased 47% between 1990 and 2010[2], and ICH risk appears to vary among white, black and Hispanic populations[3–6]. Compared to whites, young and middle-aged blacks have an almost twofold increased risk for ICH[3,4]. Similarly, Hispanics have a relative risk increase that ranges from 1.4 to 3.7 for lobar and non-lobar ICH, respectively[5]. Moreover, not only is hypertension prevalence among the elderly lower among non-Hispanic whites (76.3%) than non-Hispanic blacks (82.5%), or Hispanics (79.2%), but the risk of ICH in the presence of hypertension increases more than 50% from whites to Hispanics[7–9]. The contributions of genetic and acquired ICH risk factors to these observed risk differences are poorly understood.

Prior studies conducted in predominantly European-ancestry populations have demonstrated that Apolipoprotein E *(APOE)* ε2 and ε4 alleles potently increase risk of lobar ICH[10]. In Alzheimer's disease (AD), another disorder strongly associated with *APOE* ε allele status, the degree of risk contributed by *APOE* genotype is highly correlated with the ancestry of the population studied. Among non-Hispanic whites, homozygous carriers of *APOE* ε4 exhibit up to 12 times higher risk of AD, but this same haplotype exerts little or no effect among blacks or Hispanics[11–13].

Understanding how genetic risk factors vary across race and ethnicity may highlight novel underlying disease mechanisms and identify populations who may be particularly responsive to specific prevention strategies, as has previously been shown in treatment response for heart failure by race and ethnicity[14]. Unfortunately, with individuals of African American and Hispanic ancestry representing less than 4% of all samples in genome-wide association studies (GWAS), only recently has it become possible to study genetic risk of common disease across representative US populations[15].

We tested the effects of *APOE* ε alleles on risk of lobar and non-lobar ICH among whites, blacks, and Hispanics, using direct genotyping data supplemented by genome-wide genotyping where available in cases and controls from the International Stroke Genetics Consortium (ISGC). Because these analyses revealed substantial heterogeneity by race and ethnicity, we further explored the degree to which differential burden of hypertension across populations contributes to the variability in observed *APOE* effects.

METHODS

Participating Studies and Data Collection

Case and control subjects included in the study were gathered from 3 multicenter studies in the US and from 8 European sites participating in the ISGC, based on availability of directly ascertained *APOE* ε genotypes and a harmonized local acute case recruitment scheme. ICH cases from population-based cohorts were not included due to potential imbalances in lethal case recruitment between the two sampling approaches[16]. Studies included The Genetics of Cerebral Hemorrhage with Anticoagulation (GOCHA) study[17], the Genetic and Environmental Risk Factors for Hemorrhagic Stroke (GERFHS) study[18], the Ethnic/Racial Variations of Intracerebral Hemorrhage (ERICH) study[19], the Hospital del Mar and Vall d'Hebron Hospital ICH studies[20,21], the Jagiellonian University Hemorrhagic Stroke Study[22], the Lund Stroke Register study[23], the Edinburgh Stroke Study and LINCHPIN[24], the UMC Utrecht ICH study, and the Brescia Stroke Registry[25]. Because of variable sample sizes from contributing centers, data from European studies were analyzed together for association testing in meta-analysis (ISGC Europe), as done previously[26,27]. More specific inclusion and exclusion criteria for each of the included studies are reported in eTable 1. Demographic variables, including self-identified race and ethnicity[8] were systematically obtained from structured patient and family member interview within each site[19,28], along with additional covariates[29]. CT images on admission were analyzed at each participating site for classification as lobar (involving predominantly

the cortex and underlying white matter), and non-lobar (involving predominately the basal ganglia, periventricular white matter, or internal capsule), following prespecified criteria [26,27]. *APOE* genotype was centrally determined following standard procedures[30]. Genome-wide data were available for a subgroup of subjects. Genetic and bioinformatic analysis followed standardized prespecified quality control procedures[31] (see supplementary methods). IRB approval was obtained at all participating centers, and informed consent was obtained from all participants or their legally authorized representative.

Population Stratification

Fifteen ancestry informative markers were selected from subjects with direct or genome-wide genotyping and subjected to principal component analysis in accordance with previously published methods[32–35]. The first four principal components were included in regression models to adjust for population stratification in this subgroup. This PC analysis was not used to reclassify participants, as self-identified race/ethnicity may capture exposures that transcend genetic ancestry and could contribute to explain the stratification among different populations. A complete description of the genetic analysis, the subjects genotyped, and the markers selected is available in the supplements (eTable 2).

Statistical Analysis

Categorical variables were expressed as count (%), and continuous variables as median (interquartile range, IQR) or mean (standard deviation, SD), as appropriate. Categorical

variables were compared using the χ^2 test whereas continuous variables were compared with Mann-Whitney tests.

We tested *APOE* allele association with ICH risk using three logistic regression models. Model 1 was adjusted for age, sex, and history of hypertension. Model 2 included variables from Model 1 in addition to history of hypercholesterolemia, history of previous ischemic stroke, warfarin, statin and antiplatelet use, smoking and alcohol use. Model 3 also included variables from Model 1 and added the first four principal components (PCs) derived from ancestry-informative genotypes. *APOE* risk allele status was modeled as two variables, ε2 and ε4 coded for allele counts (0, 1, or 2 for each) in an additive model referent to the wildtype ε3 allele[17]. Analyses were performed in lobar and non-lobar ICH, given the known differences in underlying biology between the two ICH locations[36]. All statistical analyses were performed using STATA (Version 13.0; Stata Corp) and R statistical software (http://www.r-project.org).

Trans-ethnic Meta-analysis

We applied a two-stage clustering approach for meta-analysis, based on race/ethnicity and stratified by study[37]. Cases and controls in each study were divided into blacks, whites, and Hispanics, based on self-identified race/ethnicity. Each race/ethnicity group within each study were allocated to the same cluster and tested using regression models described above. Individual cluster results were presented graphically by plotting odds ratio estimates on a Forest plot to visually assess heterogeneity. The effects sizes obtained were then used

for a Der Simonian-Laird random effects, inverse-weighted non-parametric meta-analysis[38]. Cochrane's Q and I^2 were used to quantify heterogeneity.

Propensity Score Modeling of *APOE* and Hypertension

To address imbalances in the burden of hypertension across ICH populations, and related imbalances of baseline characteristics among subjects with and without hypertension, two propensity score (PS) analyses were performed using the nearest neighbor matching method to compare subjects of similar underlying hypertensive pathophysiology burden[39,40]. The first PS analysis was constructed based on history of hypertension, and included variables of age, sex, and self-identified ethnicity/race. The second PS analysis, leveraging data only available in the ERICH study, contained the same variables as the first PS analysis, in addition to the number of medications prescribed to treat hypertension, and systolic and diastolic blood pressure at ICH presentation. Propensity score results were used in a logistic regression model for ICH risk identical to Model 1 described above. As a sensitivity analysis, the same propensity score procedure was tested against age (greater than or less than 65 years), sex, and hypercholesterolemia, to increase the confidence that the PS findings were specific for hypertension.

Power Calculation

Using empiric data from our analyses, we performed a post-hoc calculation of our statistical power to detect an effect of *APOE* ε alleles on lobar ICH risk in blacks and Hispanics commensurate with the effect size detected in whites. Type I error rate of 0.05,

log additive inheritance mode, and 0.01 of population risk were assumed, with analyses

performed using Quanto software version 1.2.4 (http://hydra.usc.edu/gxe)[41].

RESULTS

13,124 subjects (47.2% cases) were included from the participating studies. 54.5% were males, median age was 66 years (inter-quartile range (IQR): 56-76), with 8,334 whites, 2,273 blacks, 1,781 Hispanics, and 736 subjects of other race/ethnicity (Table 1). The latter were excluded from the primary analyses given low statistical power. Rates of *APOE* ε4 homozygosity in cases were 3.6%, 5.3%, and 1.8%, whereas rates of *APOE* ε2 homozygosity in cases were 1.0%, 1.2%, and 0.4% respectively in whites, blacks and Hispanics. 56.8% subjects (4,069 cases and 3,379 controls) had genome-wide or direct genotyping data on ancestry informative markers for PC analysis (eTable 3). Self-identified race and ethnicity showed overall strong concordance with PC-based ancestry (eFigure 1). Additional clinical covariates were available for a subset of subjects (eTable 4).

Lobar ICH

We analyzed 2,305 lobar ICH cases from all studies. Model 1 confirmed the previously reported effect of *APOE* ε2 (pooled odd ratio (OR) = 1.49, 95 % confidence intervals (CI) = 1.24-1.80, p <0.001) and *APOE* ε4 (pooled OR = 1.51, 95% CI =1.23-1.85, p < 0.001) on ICH risk, however within self-identified Hispanics and blacks, no associations were found (Figure 1). Model 2 was used to interrogate the independent effect of *APOE* alleles on ICH, controlled for established ICH predictors (Figure 2). Here, *APOE* ε2 and ε4 allele retained association with lobar ICH. As with Model 1, this effect was observed in whites, but not in blacks or Hispanics (for *APOE* ε2, OR =1.45, 95% CI = 1.04-2.03, p = 0.028;

for *APOE* ε4, OR = 1.51, 95% CI = 1.14-1.99, p = 0.004). Model 3 considered population stratification (Figure 3). In whites, both *APOE* ε2 (OR = 1.81, 95% CI = 1.33-2.45, p < 0.001) and *APOE* ε4 (OR = 1.80, 95% CI = 1.33-2.44, p < 0.001) conferred higher risk for lobar ICH. For *APOE* ε4 alone we found a similar effect in Hispanics, suggesting that population stratification may have played some role in the lack of ε4 effect found in Models 1 and 2, particularly for the large and ethnically diverse Hispanic population recruited through the ERICH study. In contrast, for blacks neither *APOE* ε2 nor *APOE* ε4 conferred a significant risk for lobar ICH after controlling for population structure.

Non-lobar ICH

We analyzed 3,897 non-lobar ICH cases (Figure 1). In Model 1, *APOE* ε2 and ε4 did not show an effect on non-lobar ICH risk, across any of the self-identified race/ethnicity groups. When comparing non-lobar ICH cases vs controls, *APOE* ε4 p-values were 0.219 for whites, 0.182 for blacks, and 0.346 for Hispanics. For Model 2 and model 3 in non-lobar ICH, again neither *APOE* ε2 nor *APOE* ε4 showed an association with disease risk across all the studies and races/ethnicities (Figures 2 and 3).

Power calculation (lobar ICH):

Given the differences in sample sizes between whites, blacks, and Hispanics, we performed post-hoc power calculations to determine whether our study was powered to detect a comparable *APOE* effect in the smaller populations of blacks and Hispanics. Given the frequency of *APOE* ε4 in blacks (ε4 frequency 37.7%), our sample size (assuming an

unmatched case-control ratio of 1:1) would provide 99% power to detect an ε4 effect similar to the lower bound of the 95% CI seen in whites (OR=1.43). Our analyses of *APOE* ε4 effects in Hispanics were similarly powered at 90%. Further, the *APOE* ε2 frequency in blacks (19.9%) at the reported sample sizes would provide 93.8% power to detect the lower bound of the effect seen in whites (OR=1.38). For Hispanics, given the lower *APOE* ε2 frequency (0.8%) 80% power would be achieved at a slightly higher effect size (OR=1.60), but still below the one found in whites.

Propensity Score Modeling for Hypertension

We used a propensity score (PS) analysis to attempt to isolate the influence of *APOE* against the imbalanced burden of hypertension across race/ethnicity. In our first PS, we selected case and control subjects with a balanced hypertension burden, comprised of individuals of white, black, and Hispanic ancestry. In this matched and homogeneous group, we were able to detect an effect of *APOE* ε4 on lobar ICH risk among Hispanics (OR = 1.14, 95% CI = 1.03-1.28, p = 0.01), but not in blacks (OR = 1.02 95% CI = 0.98-1.07, p = 0.251). Results were confirmed in our secondary PS analysis performed only in the ERICH dataset, which included hypertension diagnosis as well as additional hypertension severity variables including number of medications used to treat hypertension, and systolic and diastolic blood pressure readings (Figure 4).

DISCUSSION

Although *APOE* effects on ICH risk have been characterized in multiple prior studies and meta-analyses for European, and more recently Asian, ancestries, there have been fewer opportunities for examination of US minority populations at disproportionate risk for ICH. Supplemented by data from the ERICH study, we are now able to confirm variability in associations between *APOE* ε genotypes and lobar ICH risk across whites, blacks, and Hispanics, and explore the degree to which differences in genetic risk are attributable to comorbid exposures. Our results demonstrate an effect of *APOE* ε4 and ε2 alleles in lobar ICH led primarily by white individuals and confirmed by additional models adjusting for known covariates[29]. When the effect of hypertension is propensity-matched across race and ethnicity, *APOE* ε4 emerges as a risk factor for lobar ICH among self-identified Hispanic individuals.

Our results highlight the challenges in generalizing genetic risk factors across ancestries, where non-genetic exposures are known to vary by race and ethnicity. In AD, the relative risks for Hispanics or blacks associated with an *APOE* ε 4 allele become progressively weaker or disappear entirely in comparison to whites[42–46]. In ICH, *APOE* ε alleles have already been shown to exert higher effects in East Asians when compared to subjects of European ancestry[47]. While recent analyses by Sawyer et al.[48] demonstrate the effect of hypertension and *APOE* ε allele status on ICH risk across race/ethnicity specific to the ERICH study, the present analysis benefits from a larger sample size via formal trans-

ethnic meta-analysis as well as a propensity score matching approach that helps to illustrate the potential mechanisms underlying the observed variability of *APOE* ε alleles on lobar ICH risk across populations.

It is important to note that the observed differences in association between *APOE* ε alleles and lobar ICH risk do not provide direct evidence that biological effects of the *APOE* gene or associations with underlying cerebral amyloid angiopathy (CAA, a major cause of lobar ICH) are necessarily different across racial and ethnic boundaries. It would seem more likely that genetic and/or environmental risk exposures covarying with race and ethnicity exert a role in modifying or mitigating underlying *APOE* genetic effects. Our propensity score analysis supports this conjecture, demonstrating that hypertension, the most important known risk factor for ICH, may simply obscure *APOE* effects that may indeed be common across ancestries. Aside from variation in environmental risk exposures, variants in a modifier gene (or genes) that differ across populations may alter the biological effect of APOE and consequently vary ICH risk, as has been hypothesized for AD[49,50]. Furthermore, genetic variants that are racially stratified and not related to APOE may directly modify the risk of ICH. This hypothesis represents an alternative explanation for why the propensity score matching for hypertension only partially remediated the effect of APOE ε4 on ICH risk in Hispanics, and had little to no effect in blacks. APOE interaction studies and trans-ethnic GWAS for ICH will likely provide insights on these hypotheses. Similarly, analyses of CAA in non-white populations will additionally clarify the effect of race/ethnicity on this pathological pathway.

In this study, we have not attempted to stratify subsets of subjects by probable CAA status using MRI data as has been done in prior meta-analyses. In fact, most prior studies linking lobar hemorrhage locations to the pathological diagnosis of CAA were performed in largely white populations[52]. As such, widely accepted criteria for classifying probable and possible CAA using hemorrhage location and microbleed counts have not been validated in non-white populations[53], Validating CAA burdens across multi-ethnic populations will require concomitant neuroimaging and/or tissue pathology data in genotyped individuals of many races and ethnicities to ensure patients are not mis-assigned.

Previously demonstrated associations between *APOE* ε4 and non-lobar ICH risk, also seen in non-lobar ICH recurrence[17,54], were not replicated in this study. Potential explanations include a higher rate of subjects affected by hypertension and an overall younger age of subjects in our study. These factors may reflect the driving effects of environmental or non-*APOE* genetic exposures in younger populations with non-lobar ICH in particular. Demographic heterogeneity is also higher in our study and the reduced availability of covariates such as steady-state lipid levels[54] for risk modeling may have influenced this finding. Finally, our previous meta-analysis of *APOE* effects in ICH also failed to show the association between ε4 and non-lobar ICH in blacks, a finding supported by the present analyses[17]. Future studies in larger datasets with well-phenotyped cases will be needed to further elucidate the potential role of *APOE* ε4 in non-lobar ICH.

The targeted enrollment of Hispanics and blacks through the ERICH study, lacking in prior reports[17,55], is an important strength of the present study. The high number of non-whites enrolled permits well-powered analyses in these populations and promotes confidence that the lack of observed effects is not a false acceptance of the null hypothesis, as supported by our *post hoc* power calculations. Some limitations should be acknowledged. Diagnosis of comorbidities are based on self-identified attestation and are therefore influenced by patient or caregiver awareness. This concern is present in both cases and controls, however, and internal consistency between diagnoses and prescribed medications helps to limit this potential source of bias. Furthermore, our propensity score is based on variables only partially capturing the complex phenotype represented by hypertension. However, this lack of information content is likely to bias our score results towards the null; we expect that a more precise index of hypertension burden would have increased our ability to normalize this phenotype across ethnicities and demonstrate an even more homogeneous *APOE* ε4 effects. Finally, genome-wide genotypes for ERICH participants are not currently available, preventing us from determining whether additional genetic exposures modify the role of *APOE* on ICH risk across race/ethnicity.

In this meta-analysis, *APOE* ε2 and ε4 remain genetic risk factors for lobar ICH but these results are largely driven by the strong associations in white individuals. However, our results support a biological effect of *APOE* ε4 alleles that seems to transcend ancestral backgrounds[56], albeit with varying effect due to the presence of racial and ethnic disparities

across related risk factors. As availability of genetic data on US minority populations continues to increase, it is hoped that improved modeling of covarying genetic and non-genetic exposures in these populations will provide new insights into treatment and prevention strategies in ICH that maximize the potential benefits for all.

POTENTIAL CONFLICTS OF INTEREST:

Dr. Anderson is supported by grants from the NIH, the American Heart Association, and the Massachusetts General Hospital Center for Genomic Medicine, and has consulted for ApoPharma, Inc.

Prof Klijn is supported by a clinical established investigator grant from the Dutch Heart Foundation (grant number 2012T077), and an Aspasia grant form Zonmw (grant number 045008048).

Table 1: Demographic characteristics, clinical and *APOE* allele frequencies across participating studies.

STUDIES		ERICH n=5017	GOCHA n=2297	ISGC Europe n=3471	GERFHS n=2339
Male sex, n (%)		2866 (57.1)	1266 (55.1)	1891 (54.6)	1130 (48.3)
Age, median (IQR)		61 (52-72)	73 (65-80)	70 (61-77)	65 (51-75)
Cases, n (%)		2880 (57.4)	1322 (57.6)	1281 (36.9)	811 (34.7)
Lobar ICH, n (%)		882(30.6)	613(47.8)	493(40.2)	316(39.0)
Non-lobar ICH, n (%)		1998 (69.4)	670 (52.2)	734 (59.8)	495 (61.0)
Hypertension, n (%)		3364/4976 (67.6)	1667/2275 (73.3)	1673/2893 (57.8)	1264/2337 (54.1)
Self-reported race/ethnicity, n (%)	white	1739 (34.7)	2024 (88.1)	2622 (75.5)	1949 (83.3)
	blacks	1751 (34.9)	131 (5.7)	-	390 (16.7)
	Hispanics	1527 (30.4)	60 (2.6)	194 (5.6)	-
	other/missing	-	82 (3.6)	654 (18.8)	-
APOE ε4 allele count, n (%)	0	3553 (70.8)	1664 (71.8)	2789 (80.4)	1664 (71.1)
	1	1298 (25.9)	570 (24.8)	637 (18.4)	601 (25.7)

	2	166 (3.3)	77 (3.4)	45 (1.3)	74 (3.2)
APOE ε2 allele count, n (%)	0	4262 (85.0)	1916 (83.4)	3034 (87.4)	1881 (80.4)
	1	710 (14.2)	363 (15.8)	413 (11.9)	431(18.4)
	2	45 (0.9)	18 (0.8)	24 (0.7)	27 (1.2)

FIGURE LEGENDS:

Figure 1:

Forest plots of meta-analysis of APOE in lobar and non-lobar ICH cases and controls in Model 1, stratified across participating studies and race/ethnicity.

Figure 2:

Forest plots of meta-analysis of APOE in lobar ICH and non-lobar ICH cases and controls, in Model 2 (adjusted for age, sex, history of hypertension, hypercholesterolemia, warfarin, statin and antiplatelet use, smoking and alcohol use), stratified across participating studies and race/ethnicity.

Figure 3:

Forest plots of meta-analysis of APOE in Lobar and Non-lobar ICH cases and controls, in Model 3 (adjusting for Model 1 covariates plus principal components 1 and 2), stratified across participating studies and race/ethnicity.

Figure 4:

Risk of APOE ε4 allele for lobar ICH, across different race/ethnicities after propensity score matching based on hypertension burden.

REFERENCES:

1.   An SJ, Kim TJ, Yoon B-W. Epidemiology, Risk Factors, and Clinical Features of

        Intracerebral Hemorrhage: An Update. *J Stroke*. 2017;19(1):3-10.

        doi:10.5853/jos.2016.00864.

2.   Krishnamurthi R V., Moran AE, Forouzanfar MH, et al. The global burden of

        hemorrhagic stroke: A summary of findings from the GBD 2010 study. *Glob Heart*.

        2014;9(1):101-106. doi:10.1016/j.gheart.2014.01.003.

3.   Qureshi AI, Giles WH, Croft JB. Racial differences in the incidence of intracerebral

        hemorrhage: effects of blood pressure and education. *Neurology*. 1999;52(8):1617-1621.

4.   Broderick JP, Brott T, Tomsick T, Huster G, Miller R. The risk of subarachnoid and

        intracerebral hemorrhages in blacks as compared with whites. *N Engl J Med*.

        1992;326(11):733-736. doi:10.1056/NEJM199203123261103.

5.   Labovitz DL, Halim A, Boden-Albala B, Hauser WA, Sacco RL. The incidence of deep

        and lobar intracerebral hemorrhage in whites, blacks, and Hispanics. *Neurology*.

        2005;65(4):518-522. doi:10.1212/01.wnl.0000172915.71933.00.

6.   van Asch CJJ, Luitse MJA, Rinkel GE, van der Tweel I, Algra A, Klijn CJM. Incidence,

        case fatality, and functional outcome of intracerebral haemorrhage overtime, according to

        age, sex, and ethnic origin: a systematic review and meta-analysis. *LANCET Neurol*.

        2010;9(2):167-176. doi:10.1016/S1474-4422(09)70340-0.

7.   Giles T, Aranda JM, Suh DC, et al. Ethnic/racial variations in blood pressure awareness,

treatment, and control. *J Clin Hypertens (Greenwich)*. 2007;9(5):345-354. doi:10.1111/j.1524-6175.2007.06432.x.

8.   Walsh KB, Woo D, Sekar P, et al. Untreated Hypertension: A Powerful Risk Factor for Lobar and Nonlobar Intracerebral Hemorrhage in Whites, Blacks, and Hispanics. *Circulation*. 2016;134(19):1444-1452. doi:10.1161/CIRCULATIONAHA.116.024073.

9.   Sturgeon JD, Folsom AR, Longstreth WT, Shahar E, Rosamond WD, Cushman M. Risk factors for intracerebral hemorrhage in a pooled prospective study. *Stroke*. 2007;38(10):2718-2725. doi:10.1161/STROKEAHA.107.487090.

10.   Carpenter A, Singh I, Gandhi C, Prestigiacomo C. Genetic risk factors for spontaneus intracerebral haemorrhage. *Nat Rev*. 2016;12:40-49.

11.   Reitz C, Mayeux R. Genetics of Alzheimer's disease in Caribbean Hispanic and African American Populations. *Biol Psychiatry*. 2014;75(5):534-541. doi:10.1002/nbm.3066.Non-invasive.

12.   Murrell JR, Price B, Lane KA, et al. Association of Apolipoprotein E Genotype and Alzheimer Disease in African Americans. *Arch Neurol*. 2006;63(3):431. doi:10.1001/archneur.63.3.431.

13.   Michaelson DM. APOE ε4: The most prevalent yet understudied risk factor for Alzheimer's disease. *Alzheimer's Dement*. 2014;10:861-868. doi:10.1016/j.jalz.2014.06.015.

14.   Taylor MR, Sun AY, Davis G, Fiuzat M, Liggett SB, Bristow MR. Race, common genetic variation, and therapeutic response disparities in heart failure. *JACC Hear Fail*.

2014;2(6):561-572. doi:10.1016/j.jchf.2014.06.010.

15. Popejoy AB, Fullerton SM. Genomics is failing on diversity. *Nature*.
   2016;538(7624):161-164. doi:10.1038/538161a.

16. Anderson CD, Nalls MA, Biffi A, et al. The effect of survival bias on case-control genetic
   association studies of highly lethal diseases. *Circ Cardiovasc Genet*. 2011;4(2):188-196.
   doi:10.1161/CIRCGENETICS.110.957928.

17. Biffi A, Sonni A, Anderson CD, et al. Variants at APOE influence risk of deep and lobar
   intracerebral hemorrhage. *Ann Neurol*. 2010;68(6):934-943. doi:10.1002/ana.22134.

18. Woo D, Sauerbeck LR, Kissela BM, et al. Genetic and environmental risk factors for
   intracerebral hemorrhage: Preliminary results of a population-based study. *Stroke*.
   2002;33(5):1190-1195. doi:10.1161/01.STR.0000014774.88027.22.

19. Woo D, Rosand J, Kidwell C, et al. The ethnic/racial variations of intracerebral
   hemorrhage (ERICH) study protocol. *Stroke*. 2013;44(10):1-7.
   doi:10.1161/STROKEAHA.113.002332.

20. Gomis M, Ois A, Rodríguez-Campello A, et al. Outcome of intracerebral haemorrhage
   patients pre-treated with statins. *Eur J Neurol*. 2010;17(3):443-448. doi:10.1111/j.1468-
   1331.2009.02838.x.

21. Domingues-Montanari S, Hernandez-Guillamon M, Fernandez-Cadenas I, et al. ACE
   variants and risk of intracerebral hemorrhage recurrence in amyloid angiopathy. *Neurobiol
   Aging*. 2011;32(3):551.e13-551.e22. doi:10.1016/j.neurobiolaging.2010.01.019.

22. Pera J, Slowik A, Dziedzic T, Pulyk R, Wloch D, Szczudlik A. Glutathione peroxidase 1

C593T polymorphism is associated with lobar intracerebral hemorrhage. *Cerebrovasc Dis*. 2008;25(5):445-449. doi:10.1159/000126918.

23.   Hallström B, Jönsson AC, Nerbrand C, Norrving B, Lindgren A. Stroke incidence and survival in the beginning of the 21st century in Southern Sweden: Comparisons with the late 20th century and projections into the future. *Stroke*. 2008;39(1):10-15. doi:10.1161/STROKEAHA.107.491779.

24.   Deary IJ, Gow AJ, Pattie A, Starr JM. Cohort profile: The lothian birth cohorts of 1921 and 1936. *Int J Epidemiol*. 2012;41(6):1576-1584. doi:10.1093/ije/dyr197.

25.   Pezzini A, Grassi M, Iacoviello L, et al. Serum cholesterol levels, HMG-CoA reductase inhibitors and the risk of intracerebral haemorrhage. The Multicenter Study on Cerebral Haemorrhage in Italy (MUCH-Italy). *J Neurol Neurosurg Psychiatry*. 2016;87(9):924-929. doi:10.1136/jnnp-2015-312736.

26.   Woo D, Falcone GJ, Devan WJ, et al. Meta-analysis of genome-wide association studies identifies 1q22 as a susceptibility locus for intracerebral hemorrhage. *Am J Hum Genet*. 2014;94(4):511-521. doi:10.1016/j.ajhg.2014.02.012.

27.   Anderson CD, Falcone GJ, Phuah CL, et al. Genetic variants in CETP increase risk of intracerebral hemorrhage. *Ann Neurol*. 2016;80(5):730-740. doi:10.1002/ana.24780.

28.   Marini S, Morotti A, Ayres AM, et al. Sex differences in intracerebral hemorrhage expansion and mortality. *J Neurol Sci*. 2017;379:112-116. doi:10.1016/j.jns.2017.05.057.

29.   Woo D, Kissela BM, Khoury JC, et al. Hypercholesterolemia, HMG-CoA reductase inhibitors, and risk of intracerebral hemorrhage: a case-control study. *Stroke*.

2004;35(6):1360-1364. doi:10.1161/01.STR.0000127786.16612.A4.

30. Koch W, Ehrenhaft A, Griesser K, et al. TaqMan systems for genotyping of disease-related polymorphisms present in the gene encoding apolipoprotein E. *Clin Chem Lab Med*. 2002;40(11):1123-1131. doi:10.1515/CCLM.2002.197.

31. Anderson C, Pettersson F, Clarke G, Cardon L, Morris A, Zondervan K. Data quality control in genetic case-control association studies. *Nat Protoc*. 2010;5(9):1564-1573. doi:10.1038/nprot.2010.116.

32. Purcell S, Neale B, Todd-Brown K, et al. PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet*. 2007;81(3):559-575. doi:10.1086/519795.

33. Biffi A. Genetic Variation and Neuroimaging Measures in Alzheimer Disease. *Arch Neurol*. 2010;67(6):677. doi:10.1001/archneurol.2010.108.

34. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick N a, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet*. 2006;38(8):904-909. doi:10.1038/ng1847.

35. Marini S, Lena UK, Crawford KM, et al. Comparison of Genetic and Self-Identified Ancestry in Modeling Intracerebral Hemorrhage Risk   . *Front Neurol* . 2018;9:514. https://www.frontiersin.org/article/10.3389/fneur.2018.00514.

36. Devan W, Falcone G, Anderson C, et al. Heritability Estimates Identify a Substantial Genetic Contribution to Risk and Outcome of Intracerebral Hemorrhage. *Stroke*. 2014;44(6):1578-1583. doi:10.1016/j.pestbp.2011.02.012.Investigations.

37. Hong J, Lunetta KL, Cupples LA, Dupuis J, Liu CT. Evaluation of a Two-Stage Approach in Trans-Ethnic Meta-Analysis in Genome-Wide Association Studies. *Genet Epidemiol*. 2016;40(4):284-292. doi:10.1002/gepi.21963.

38. DerSimonian R, Laird N. Meta-analysis in clinical trials revisited. *Contemp Clin Trials*. 2015;45:139-145. doi:10.1016/j.cct.2015.09.002.

39. Austin PC. An introduction to propensity score methods for reducing the effects of confounding in observational studies. *Multivariate Behav Res*. 2011;46(3):399-424. doi:10.1080/00273171.2011.568786.

40. Ho D, Imai K, King G, Stuart E. Matching as Nonparametric Preprocessing for Reducing Model Dependence in Parametric Causal Inference. *Polit Anal*. 2007;15:199{\textendash}236.

41. Evans DM, Purcell S. Power calculations in genetic studies. *Cold Spring Harb Protoc*. 2012;7(6):664-674. doi:10.1101/pdb.top069559.

42. Tang MX, Stern Y, Marder K, et al. The APOE-epsilon4 allele and the risk of Alzheimer disease among African Americans, whites, and Hispanics. *JAMA*. 1998;279(10):751-755. doi:10.1001/jama.279.10.751.

43. O'Bryant SE, Johnson L, Reisch J, et al. Risk factors for mild cognitive impairment among Mexican Americans. *Alzheimer's Dement*. 2013;9(6):622-631.e1. doi:10.1016/j.jalz.2012.12.007.

44. Farrer LA, Cupples LA, Haines JL, et al. Effects of age, sex, and ethnicity on the association between apolipoprotein E genotype and Alzheimer disease. A meta-analysis.

APOE and Alzheimer Disease Meta Analysis Consortium. *JAMA*. 1997;278(16):1349-1356. doi:10.1001/jama.1997.03550160069041.

45.   Evans DA, Bennett DA, Wilson RS, et al. Incidence of Alzheimer disease in a biracial urban community: Relation to apolipoprotein E allele status. *Arch Neurol*. 2003;60(2):185-189. doi:10.1001/archneur.60.2.185.

46.   Graff-Radford NR, Green RC, Go RCP, et al. Association between apolipoprotein E genotype and Alzheimer disease in African American subjects. *Arch Neurol*. 2002;59(4):594-600. doi:10.1001/archneur.59.4.594.

47.   Tzourio C, Arima H, Harrap S, et al. APOE genotype, ethnicity, and the risk of cerebral hemorrhage. *Neurology*. 2008;70(16):1322-1328. doi:10.1212/01.wnl.0000308819.43401.87.

48.   Sawyer R, Sekar P, Osborne J, et al. Racial/ethnic variation of APOE alleles for lobar intracerebral hemorrhage. *Neurology*. 2018;91(5).

49.   Lee JH, Cheng R, Vardarajan B, et al. Genetic modifiers of age at onset in carriers of the G206A mutation in PSEN1 with familial Alzheimer disease among Caribbean hispanics. *JAMA Neurol*. 2015. doi:10.1001/jamaneurol.2015.1424.

50.   Maestre G, Ottman R, Stern Y, et al. Apolipoprotein E and alzheimer's disease: Ethnic variation in genotypic risks. *Ann Neurol*. 1995. doi:10.1002/ana.410370217.

51.   Williams FMK, Carter AM, Hysi PG, et al. Ischemic stroke is associated with the ABO locus: the EuroCLOT study. *Ann Neurol*. 2013. doi:10.1002/ana.23838.

52.   Charidimou A, Martinez-Ramirez S, Reijmer YD, et al. Total Magnetic Resonance

Imaging Burden of Small Vessel Disease in Cerebral Amyloid Angiopathy. *JAMA Neurol*. 2016;73(8):994. doi:10.1001/jamaneurol.2016.0832.

53.   Knudsen KA, Rosand J, Karluk D, Greenberg SM. Clinical diagnosis of cerebral amyloid angiopathy: validation of the Boston criteria. *Curr Atheroscler Rep*. 2003;5(4):260-266. doi:10.1007/s11883-003-0048-4.

54.   Raffeld MR, Biffi A, Battey TW, et al. APOE epsilon4 and lipid levels affect risk of recurrent nonlobar intracerebral hemorrhage. *Neurology*. 2015;85(4):349-356. doi:10.1212/wnl.0000000000001790.

55.   Woo D, Kaushal R, Chakraborty R, et al. Association of apolipoprotein E4 and haplotypes of the apolipoprotein E gene with lobar intracerebral hemorrhage. *Stroke*. 2005;36(9):1874-1879. doi:10.1161/01.STR.0000177891.15082.b9.

56.   Tai LM, Thomas R, Marottoli FM, et al. The role of APOE in cerebrovascular dysfunction. *Acta Neuropathol*. 2016;131(5):709-723. doi:10.1007/s00401-016-1547-z.